Age and Ageing 1997; **26**: 45–52

REVIEW

Dispersion measures in biomedical research on ageing: nuances in the meaning of variability

MARCEL G. M. OLDE RIKKERT, MARTIN A. VAN'T HOF[1], WILLIBRORD H. L. HOEFNAGELS

Department of Geriatric Medicine, University Hospital Nijmegen, St Radboud, The Netherlands
[1]Department of Methodology, Statistics and Informatics, University of Nijmegen, PO Box 9101, 6500 HB Nijmegen, The Netherlands

Address correspondence to: M. G. M. Olde Rikkert

Keywords: *dispersion measures, published empirical studies, within-group variability*

Introduction

The existence of an increase in differences between subjects for many biological, sociological and psychological parameters with age is a widespread belief in gerontology [1–6]. Both theoretical arguments and empirical studies support this hypothesis [1, 7]. By analogy with the increase in variability between age groups in many psychological and biological parameters, a similar increase in variability in biomedical and pathophysiological parameters with age is commonly assumed.

There are solid reasons for such an assumption. Firstly, there is the fact that atypical presentation of illnesses occurs very frequently in elderly people, a widely accepted principle in geriatric medicine. This individualization in presenting symptoms should result in a greater within-group variability in symptoms in older age groups compared with younger patients with the same disease. Secondly, just as ageing occurs at individual rates, the progression of degenerative illnesses may be highly individual. The highly individual nature that this implies for degenerative diseases has been described for Alzheimer's disease [8]. Finally, geriatric patients characteristically suffer from multiple diseases and impairments at the same time—and each individual will show a different combination and severity of problems. In their famous article on human ageing, Rowe and Kahn stress the importance of a differential approach to the diversity of the ageing population [9]. It seems logical to apply such a differential approach also to research on geriatric patients.

However, the effect of age on variability within age-groups is much less studied in medical gerontology than in the sociological and psychological domains of gerontology. In the present paper we report the level of attention paid to dispersion measures in recently published geriatric studies. We also studied whether sufficient data were reported to test the generally accepted age-related increase in variability between subjects in the biomedical outcome parameters studied.

Methods

Many terms are used in the description of the distribution width of a data set. However, although they are very important in statistical analysis, there are no clear definitions of commonly used terms such as heterogeneity, diversity, dispersion, variability, variation and spread. 'Heterogeneity', 'diversity', 'variation' and 'spread' will not be used in the present report because they are often used in everyday language, and in other contexts. In this study 'dispersion' refers to the distribution width of a parameter; 'dispersion measures' (DMs) are defined as all possible measures that can be used to quantify the width of a distribution. For numerical data, the most important DMs are: standard deviations (SDs)—and variances; coefficients of variation (CVs); ranges; and percentiles. The term 'variability' can be used unequivocally in statistics only as a part of a compound. It can be defined as differences between measurements carried out at the same time or change in measurements with time. In biomedical research, variability can generally be divided into a biological, an analytical (measurement error) and a temporal component. Intra-individual variability, between-group variability and within-group variability (WGV) are essentially different. In this study,

only WGV will be discussed. Homeoscedasticity and heteroscedasticity are the statistical terms that describe equality and unequality in dispersion, respectively.

Article selection

The four journals with the highest 'impact factor' in the subject category listing Geriatrics & Gerontology of the 1993 *Citation Index* were selected for this study [10]. These journals are: *Journal of the American Geriatrics Society*, *Journal of Gerontology*, *Mechanisms of Ageing and Development* and *Age and Ageing*. From the *Journal of Gerontology* only the Medical Sciences section was included. We reviewed volumes 1993 and 1994 of these journals and selected all articles presenting empirical data. Animal and case studies were excluded.

For each article we recorded: (1) the type of research; (2) the number of subjects, their mean age and, if present, standard deviations of age and age ranges; (3) the type and number of age-dependent outcome variables; (4) the presence, type and discussion of DMs; (5) the presence of a discussion of WGV; (6) the presence of the keyword 'variability' in the title or abstract.

Statistical analysis of WGV

The *F*-test of variance ratios (F_{var}: ratio of older/younger variance) was selected to test WGV between age groups. Data required for this *F*-test are: the number of younger and elderly subjects and the SD. Standard errors of the mean (SEM) or confidence intervals (CI) could also be used after transformation to SD. The parametric *F*-test was chosen despite its only moderate power as without the original data, there is no better way of testing WGV. WGV can be confounded by differences in variable means for different age groups. For this reason, an analogous *F*-test for the ratios of the squared CVs (F_{cv}) for the older and younger group was performed. A two-sided significance level of 0.05 was chosen for these *F*-tests. As well as judging each individual study for the possibility of performing these *F*-tests, we aimed to produce an overall analysis of WGV. This analysis was executed by the so-called vote-counting technique: summing significant and nonsignificant *F*-ratios within age groups [11]. In the case of an interventional study, *F*-tests were performed on baseline data and post-intervention data of the outcome variables. Only patient characteristics that were not directly related to the outcome variables, such as anthropometric measures, were not included in the analysis of WGV.

Additionally, a weighted analysis of WGV was performed by giving greater weight to those studies with larger sample size. This method was also applied by Devolder in her meta-analysis of WGV in memory functions [12]. The weight (w) of individual studies' *F*-ratios was calculated by comparing the number of younger (n_1) and elderly (n_2) subjects of an individual study with the total number of subjects (N) in all selected studies according to the formula

$$w = [n_1 \times n_2/(n_1 + n_2)]/N.$$

Each *F*-ratio was then multiplied by this weight and so, after summation, weighted overall *F*-ratios for variances and *CV* could be obtained ($F_{w,var}$ and $F_{w,cv}$ respectively).

Criteria for inclusion

Differences in WGV in biomedical variables between different age groups could only be analysed in this way if the studies fulfilled the criteria presented in Table 1. *F*-tests could only be performed in cross-sectional studies since in longitudinal studies the subsequent measurements in the same individuals at different ages are not independent and the original data would be necessary to test an increase in WGV with age. Studies were also excluded when only psychological, sociological or demographic outcome variables were measured because we wanted to focus on biomedical or functional performance variables. Functional performance variables were defined as variables measuring performance at subject level (mobility, continence, cognition and performance of activities of daily living). This type of variable was included because of the importance of functional assessment in geriatrics.

There were a number of statistical prerequisites to allow application of *F*-tests. Firstly, *F*-tests require the presence of DMs (SD, SEM or CI) of normally distributed data. To be meaningful, these DMs should address data measured quantitatively rather than on ordinal or nominal measurement scales. Furthermore, the DMs has to be present for age groups older and younger than 65 years. To exclude the confounding of WGV in the younger subjects by growth and development, the youngest subject had to be older than 20 years. A larger WGV in age in one of the groups (mostly the older age group) may be a cause of differences in WGV in the age-dependent variables. For this reason *F*-tests were not performed in cases where there were large differences in WGV in age (arbitrarily a factor of 2 was chosen). Studies were also excluded if a substantial ceiling effect in the age-dependent variables was likely. This effect may occur especially in scales for functional assessment. The WGV for the younger age group is likely to be much smaller in these cases because most of the younger subjects have maximum scores.

Results

There were 586 articles in which empirical data were presented. Only 76 studies (13%) focussed on elderly patients under the care of geriatricians. Mostly, patients

Table 1. Criteria used to judge whether age-related differences in within-group variability in recently published biomedical outcome variables could be tested with *F*-ratio tests

Presence of empirical data
Cross-sectional study design
Inclusion of outcome variables other than psychological/ sociological/demographic ones
Presence of within-group variability for different age groups (i.e. SD, SEM, CI)
Numerical data available
Mean age of elderly subjects > 65 years
Minimum age of younger subjects > 20 years, mean age < 65 years
Number of subjects > 5
Standard deviations of ages or age-ranges not more than twice as large for elderly or younger subjects
No ceiling effect in outcome variables
Presence of details of the data: normal distribution, quantitatively measured

SD, standard deviation; SEM, standard error of the mean; CI, confidence interval.

Table 2. Characterization of the 586 biomedical studies on ageing included in this review

Subjects/design	No. of studies (%)
Geriatric inpatients	43 (7)
Geriatric outpatients	33 (6)
Other disciplines	161 (27)
Nursing home patients	82 (14)
Healthy elderly people	171 (29)
Other	96 (16)
Longitudinal	111 (19)
Cross-sectional	438 (75)
Other	37 (6)
Observational	459 (78)
Intervention	127 (22)

from other disciplines (27%) or healthy or community dwelling elderly (29%) were studied (Table 2). The designs of most studies were cross-sectional and observational. Functional, psychological and sociological variables were studied in addition to biomedical variables. Of the 586 studies, 366 reported DMs quantitatively, rather than just graphically (Table 3). The study sample was divided into studies that predominantly studied biomedical, functional and psychological or sociological topics. Studies of functional performance measures presented DM more often than psychological and sociological studies, although there were relatively few (47) in the latter category. DM frequencies were similar.

More than 90% of the studies that presented some kind of DM used SD. In psychological and sociological studies, as well as in studies focussing on functional performances, it is questionable whether all variables meet the precondition of being measured on interval or ratio scales. However, the common application of SD was never justified or discussed with regard to this precondition in any of the 586 studies. Similarly, normality (which is required for meaningful interpretation of SD), was only mentioned in 13 studies (2%). Percentiles, ranges and CVs were reported less frequently. SDs and ranges were used in 15% of the whole study sample, while just 2% used SD as well as percentiles. Only five studies used box-and-whisker plots which, along with graphical frequency distributions of dependent outcome variables in the subjects (used in 18 studies), show the dispersion of data particularly well.

DMs were discussed 30 times, but WGV in different age groups was discussed separately only six times. Campbell and co-workers used *F*-ratios to analyse age differences in WGV in risk factors of coronary heart disease between two groups aged over 70 [13]. Remarkably, when two groups without coronary heart disease were compared, the WGV was smaller for glucose and body mass index in subjects of 80 years and older than in subjects of 70–79 years of age. No variables showed an increase in variability with age. The authors suggest that this narrower range may reflect better maintenance of homeostasis in the oldest subjects, which may be a factor in improved survival. Whisler and Grants found a larger WGV in the functioning of human B-lymphocytes in elderly people [14]. In contrast to these two studies were the remaining four studies that met our criteria for further analysis of WGV. Hausdorff and co-workers focussed on measures of intra-individual walking rate variability [15]. They describe a larger WGV in these measures in the case of congestive heart failure, but did not find an age-related difference in WGV. The groups of King and Baloh both discuss a larger WGV in older age in measures of posture and sway [16, 17]. Moschner and Baloh measured the effect of age on eye movements [18]. They found a larger WGV in the elderly in some, but not all variables studied. Three of the six studies discussing WGV measured aspects of posture and mobility.

The keyword 'variability' was mentioned four times in the title of this study sample. In none of these was it clear from the title what kind of variability had been studied. Twice it transpired that it was intra-individual variability [19, 20]. Bearden and co-workers used 'variability' in their title to mean the between-group differences in the mean number of performed cardiovascular imaging procedures [21]. Hausdorff *et al.* probably used 'walking variability' in the title of their article to mean both intra-individual variability and WGV [15]. Another 15 studies used this keyword in the abstract. Three times it was used to mean intra-individual variability, twice between-group variability,

Table 3. Frequency distribution of the usage of the most important dispersion measures (DMs) for all included empirical studies, and for a subdivision in three different types of outcome variables

	No. (and %) of studies			
DMs inclusion	All ($n = 586$)	Biomedical ($n = 467$)	Functional ($n = 72$)	Psychological/sociological ($n = 47$)
No. of DMs				
$\geqslant 1$	366 (62)	292 (63)	55 (76)	19 (40)
0[a]	220 (38)	175 (37)	17 (24)	28 (60)
Type of DMs				
SD	338 (58)	276 (59)	45 (63)	17 (36)
Range	122 (21)	98 (21)	19 (26)	5 (11)
Percentiles	22 (4)	16 (3)	5 (7)	1 (2)
CV	6 (1)	6 (1)	–	–
Other	5 (1)	4 (1)	1 (1)	–

[a] No DMs presented numerically.
SD, standard deviation; CV, coefficient of variation.

eight times WGV, in the final two articles it had still another meaning. Only three times was the type of variability defined by the authors. In the other 12 studies what was meant had to be determined by the reader from the context.

Age-effect on WGV

In 72 studies, comparisons in biomedical or functional variables were made between groups of subjects of 65 years and over and subjects younger than 65 years in such a way that the first eight criteria for testing WGV were fulfilled. From these studies, 16 (23%) were excluded because WGV in age was unclear or more than twice as large in the elderly groups. Only one study measuring neuropsychological variables was excluded because of ceiling effects. The resulting 55 studies reported on a wide range of topics within the field of medical gerontology (Table 4). Only two studies stated explicitly that data were normally distributed. To prevent exclusion of nearly all the studies, we assumed that this condition was fulfilled by all 55. The mean number of elderly subjects per study was 31 (range 5–216). The mean number of younger subjects was 26 (range 5–211). The mean ages of elderly and younger subjects were 75 and 30, respectively. The mean number of age-dependent outcome variables per study to be tested with *F*-tests was 9.6 (sum: 527; SD: 9.1, range: 1–46).

Vote-counting was performed for the *F*-ratios of all included 527 variables. In most variables (54%) there was no significant difference in WGV between the elderly and the younger subjects. In the majority of the variables (75%) there was concordance of F_{var}- and F_{cv}-tests (Table 5). In total, SDs were significantly larger for 119 variables in the elderly. The elderly group's CVs were significantly larger, smaller or not significantly different for 126, 51 and 350 variables, respectively. Thus, vote-counting indicated that there was a considerable between-variable variability in the effect of age on WGV. The medians of F_{var} and F_{cv} were 1.30 and 1.33 respectively. The distributions of F_{var} and F_{cv} were both skewed to the right with skewnesses of 11.8 and 20.3, respectively (Figure 1). The 95th-percentiles of F_{var} and F_{cv} were 10.7 and 19.6, respectively. Weighted overall *F*-ratios were: $F_{w,var} = 4.29$; $F_{w,cv} = 7.52$. These values are considerably larger than the medians of the unweighed *F*-ratios. While there is no statistical test available for evaluating the significance of these weighted *F*-ratios, the critical *F*-value of 2.98 applied in vote-counting (in a case where the young and elderly groups had equal sample sizes of 10 subjects) could serve as a reference point.

Vote-counting was also performed for each individual study. Of the 55 studies 36% did not show a larger CV in WGV for any of the tested variables in the elderly, and 22% showed a larger CV in more than 50%

Table 4. Topics of the final study sample of 55 biomedical studies on ageing selected for meta-analysis of within-group variability

Topic	No. of studies
Endocrinology	11
Immunology	9
Mobility/posture	7
Functional performance	5
Nutrition/body composition	5
Exercise	4
Haematology	4
Neurology	2
Dermatology	2
Cardiovascular	1
Gastroenterology	1
Infectious diseases	1
Pathology	1
Oncology	1
Pharmacology	1

Table 5. 'Vote counting' for significance of 527 two-sided F-ratio tests of variances (F_{var}) and coefficients of variation (F_{cv}, $\alpha = 0.05$)

	No. (and %) of F-ratio tests			
	F_{var} old > F_{var} young	F_{var} old < F_{var} young	F_{var} non significant	Total
F_{cv} old > F_{cv} young	81 (15)	10 (2)	35 (7)	126 (24)
F_{cv} old < F_{cv} young	3 (1)	29 (6)	19 (4)	51 (10)
F_{cv} non significant	35 (7)	31 (6)	284 (54)	350 (66)
Total	119 (23)	70 (13)	338 (64)	527 (100)

of the tested variables. A higher proportion of the studies (71%) did not show a smaller CV for the elderly in any of the variables tested, and only 4% of the studies showed a significantly smaller CV for at least 50% of the variables tested. Hence, although there was a considerable between-study variability, a larger CV was found for the elderly more often than a smaller CV.

Discussion

This study is the first review on dispersion measures in which only biomedical journals on ageing were selected. Moreover, our study included more articles than earlier reviews [7, 12, 22, 23]. Most (62%) of the study sample presented DMs. This is a much larger proportion than the 43% found by Nelson and Dannefer [7], possibly because far more biomedical studies were included in our study (80% vs 44%). The 40% of psychological and sociological studies that used at least one type of DMs is more consistent with their data. However, the overall percentage of studies in which DMs findings were also the subject of discussion is even smaller than that found by these authors (7% vs 27%). Although the type of distribution and the type of measurement scale was taken into account only a few times, the SD is used as DMs in most studies. The SD is the most accurate DMs in the case of normality, but percentiles (and, graphically, box-and-whisker plots) are more informative in skewed distributions. In elderly populations non-normality might be common. However, percentiles or box-and-whisker plots were presented in addition to the SD in only nine and five studies, respectively. Studies on functional performances may frequently apply research instruments not fulfilling the criterion of using interval or ratio scales

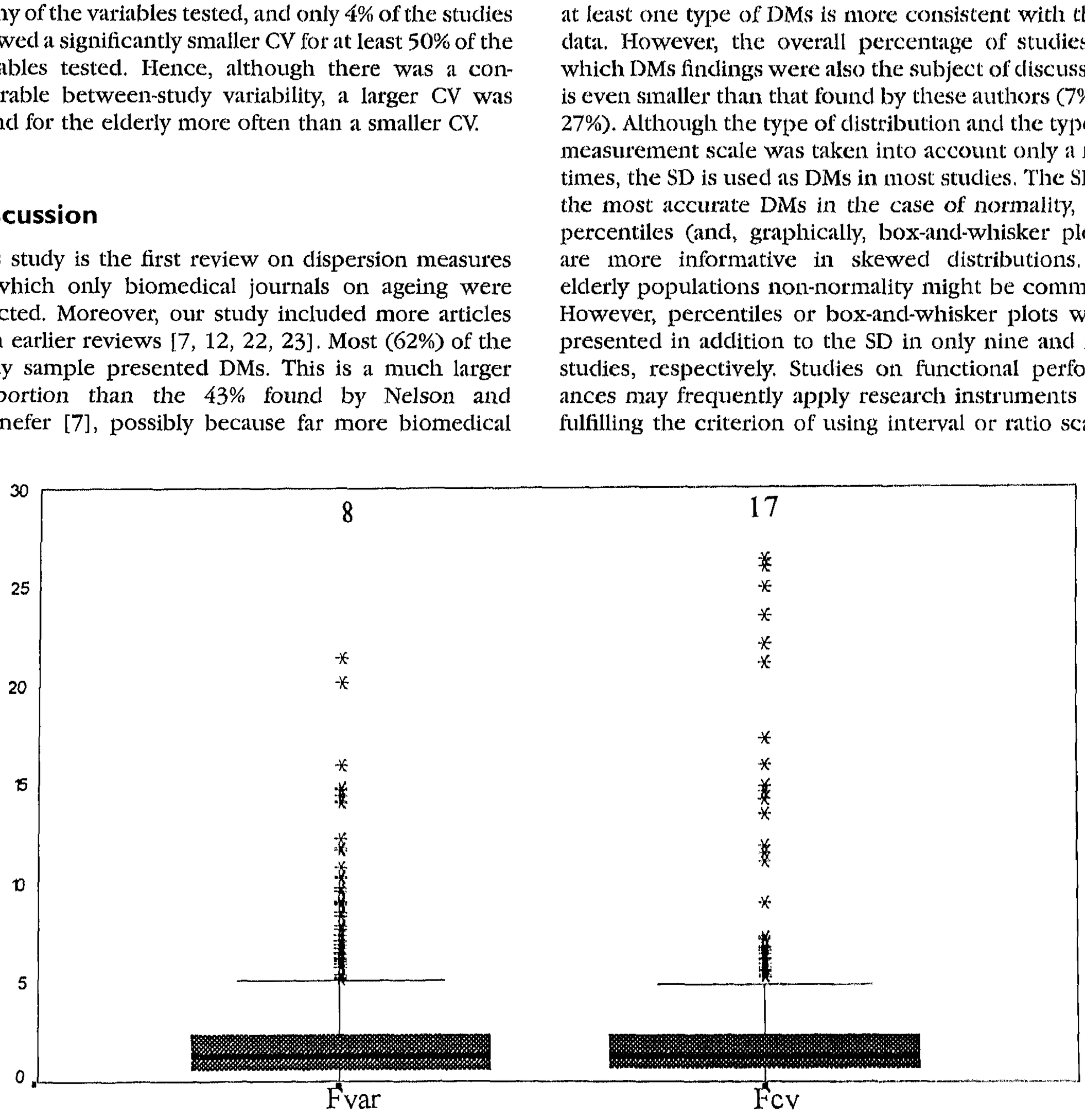

Figure 1. F-ratios (elderly subjects/younger subjects) of variances (F_{var}) and coefficients of variation (F_{cv}) of 527 biomedical variables from 55 studies on ageing. The box-plot shows: median 25th and 75th percentile and some of the outliers, more than 1.5 box-length from the upper edge of the box (*).

(e.g. Mini-Mental State Examination Score, Barthel index). However, these type of studies mostly used SD. DMs probably more suitable for data with ordinal or nominal scales (such as the index of qualitative variation) were not mentioned. Despite the frequent presentation of DMs, few authors used the presented DMs in the discussion or conclusions of the article.

Age-effect on WGV

Although there is substantial evidence to support the hypothesis of age-related increase in WGV in gerontology, it remains a topic of scientific debate. Four meta-analyses on WGV in gerontology have been published so far. Krauss, Devolder and Bornstein and Smircina examined 20, 22 and 56 articles respectively [12, 22, 23]. Kraus supplies evidence for an increase in WGV with age in multiple cognitive variables, while the other two studies question a general increase in WGV with age in the fields of memory, human behaviour and performance. The most recent and extensive review was by Nelson and Dannefer [7]. In their review they studied 127 gerontological studies published between 1982 and 1987 in six journals that focus on ageing and development. Less than half (44%) of these studies reported on biological variables. From the longitudinal and cross-sectional studies that did report these measures, 83 and 63% respectively showed an increase in the WGV for the variables that were studied. However, by studying the relation between age and WGV in reaction time, Hale *et al.* showed that the increase in WGV was the result of the slowing of reaction time with age [24]. Such an increase in the mean might also be the explanation for increase in WGV in other variables. In short, there is evidence for an increase in WGV for many cognitive, behavioural and biological variables, but for other variables such an increase is questionable. There is limited empirical support to regard an increase in WGV as a universal consequence of ageing *per se*.

In this study most of the 527 biomedical variables of the final study sample were studied in healthy or normal elderly people. Only a few studies compared clinical characteristics of elderly patients with those of younger patients with the same disease, and none of them described WGV of symptoms and signs quantitatively. Our findings give more support to a differential effect of ageing on WGV than to an overall increase of WGV with age for all kind of biomedical variables. This suggests that trying to study the overall effect of age on WGV in biomedical variables is of limited relevance. WGV seems to be more relevant for some studies and some variables than for others in comparing age groups. WGV may be most clearly increased by age in complex outcome variables such as mobility and posture than depend on one or more complex biological systems. For less complex variables, variability may or may not increase with age. The pituitary–adrenal glucocorticoid response is an example of a less complex variable in which WGV showed no increase with age [25]. However, the immunological process of B-lymphocyte stimulation, also not a variable at subject level, showed a larger WGV in the elderly [14]. Although *F*-ratios were largest for variables concerning stability, mobility and posture, or study sample is too small and the between-study variability too large to draw conclusions on how this differential ageing affects WGV in different fields of medical gerontology. In studies with large WGV it is necessary to determine the factors that contribute to it. This search may generate valuable information about the ageing process itself and factors contributing to successful ageing.

The reported age-related differences in WGV in this review may have several explanations. Firstly, differences in within-subject variability between age groups may influence findings of WGV in cross-sectional studies. The effect of an increase in within-subject variability on WGV is mentioned explicitly by Moschner and Baloh [18], but may also be present in other studies. Secondly, WGV is highly dependent on subject selection. Most studies used clinical criteria such as medical history, physical examination, drug use and laboratory tests to rule out clinical and subclinical diseases. Six studies used the so-called SENIEUR protocol, which consists of a large set of predetermined criteria to identify healthy elderly people for immunological studies [26]. In gerontological studies on exercise, strict criteria to include only healthy elderly people are developed in a similar way [27]. Both sets of criteria were developed to study the effect of age on the mean values of immunological and exercise variables for groups of subjects. However, strict criteria should also be used to study the effect of age *per se* on WGV. In our study sample, this was only done by Whisler and Grants in their study on age-related changes in functional performance of B-lymphocytes [14]. In contrast to such a careful exclusion of diseases, it seems probable that the general notion of increase in WGV with age is continuously fed by clinical observations of geriatric patients in the absence of exclusion criteria. Thirdly, ageing itself may have a differential effect on WGV in different variables. Whisler and Grants support this view, suggesting that the 'heterogeneity of the human ageing process may possibly extend to molecular mediators and events' [14]. Finally, mortality may differentially affect the elderly populations studied. Different biomedical variables have different weight in the overall mortality risk of subjects. Hence, surviving may have a complex impact on WGV of study populations. WGV of serious risk factors may be decreased by surviving, as suggested by Campbell *et al.* [13], while mortality may not affect WGV in variables that are not risk factors.

Some caution is needed in interpreting the results of the *F*-tests presented in this study. It is likely that in

many cases the parametric *F*-test will have been used for variables with a non-normal distribution. This may result in an overestimation of WGV in some studies. On the other hand, an underestimation of the number of variables with a larger WGV for the elderly is possible, because of the poor power of the *F*-tests in the case of small study samples. Because the five studies focussing on functional performances mainly measured data on ratio scales (e.g. time, speed, distance) it is not very likely that the results are seriously troubled by incorrect application of *F*-tests on data from ordinal measurement scales. However, the technique of vote-counting is possibly confounded by interdependence of measures within each study. This interdependence could not be excluded, although only variables labelled as important outcome measurements were studied.

Conclusions

This study stresses the importance of reporting and discussing dispersion measures when publishing biomedical research on ageing. The quality of the descriptive statistics probably may be improved by carefully taking into account the type of measurement scales and the skewness or normality of the data. The term 'variability' should be used more precisely by adding explanatory terms. The hypothesis of a general increase in WGV with age in all biomedical variables is not supported by this review. However, clinical data on the topic of WGV and their discussion were sparse, even though only biomedical journals on ageing were selected. Without scientific evidence, individualized alternatives for the diagnostic and therapeutic management of elderly patients are advocated in the determination of biochemical reference ranges [28], in pharmacotherapy [29] and in the treatment of geropsychiatric disorders [30]. More data are needed to ensure than an expensive request for individualizing health care for the elderly is well-founded. In conclusion, nuances in the meaning of variability must be clarified in future research on ageing.

Key points

- Published clinical data on within-group variability with age in biomedical variables are sparse.
- The term 'variability' should be used more precisely and nuances in its meaning clarified.

Acknowledgements

This research was supported by The Netherlands Programme for Research on Ageing (NESTOR), funded by the Ministry of Education, Culture and Sciences, and the Ministry of Health, Welfare and Sports.

References

1. Sprott RL. Age-related variability. Ann N Y Acad Sci 1988; 515: 121–3.

2. Maddox GL. Aging differently. Gerontologist 1987; 27: 557–64.

3. Dannefer D. Differential gerontology and the stratified life course: conceptual and methodological issues. Annu Rev Gerontol Geriatr 1988; 8: 3–36.

4. Berkman LF. The changing and heterogeneous nature of aging and longevity: a social and biomedical perspective. Annu Rev Gerontol Geriatr 1988; 8: 37–68.

5. Pesch HJ. Der Altersknochen als Paradigma für des Alterns. Z. Gerontol 1990; 23: 128–9.

6. Rapp, PR, Amaral DG. Individual differences in the cognitive and neurobiological consequences of normal aging. Trends Neurosci 1992; 15: 340–5.

7. Nelson EA, Dannefer D. Aged heterogeneity: fact or fiction? The fate of diversity in gerontological research. Gerontologist 1992; 32: 17–23.

8. Rowe JW, Kahn RL. Human aging: usual and successful. Science 1987; 237: 143–9.

9. Boller F, Forette F, Khachaturian Z, Poncet M, Christen Y (eds). Heterogeneity of Alzheimer's Disease. Berlin: Springer-Verlag, 1992.

10. Garfield E. SCI Journal Citation Reports: a Bibliometric Analysis of Science Journals in ISI database. Philadelphia: Institute for Scientific Information, 1994; 81.

11. Hunter JE, Schmidt FL, Jackson GB. Meta-analysis. Cumulating research findings across studies. Beverly Hills: Sage, 1982; 130

12. Devolder PA. Comparison of within-group variabilities across older and younger adults: a meta-analysis. Exp Aging Res 1991; 17: 113–7.

13. Campbell AJ, Busby WJ, Robertson MC. Over 80 years and no evidence of coronary heart disease: characteristics of a survivor group. J Am Geriatr Soc 1993; 41: 1333–8.

14. Whisler RL, Grants IS. Age-related alterations in the activation and expression of phosphotyrosine and protein kinase C (PKC) among human B cells. Mech Ageing Dev 1993; 71: 31–46.

15. Hausdorff JM, Forman DE, Ladin Z *et al.* Increased walking variability in elderly patients with congestive heart failure. J Am Geriatr Soc 1994; 42: 1056–61.

16. King MB, Judge JO, Wolfson L. Functional base of support decreases with age. J Gerontol 1994; 49: M258–63.

17. Baloh RW, Fife TD, Zwerling L *et al.* Comparison of static and dynamic posturography in young and older normal people. J Am Geriatr Soc 1994; 42: 405–12.

18. Moschner C, Baloh RW. Age-related changes in visual tracking. J Gerontol 1994; 48: M235–8.

19. Luutonen S, Antila K, Neuvonen P *et al.* Spectral analysis of heart rate variability in evaluation of sympathetic function in elderly subjects. Age Ageing 1994; 23: 473–7.

20. Resnick NM, Beckett LA, Branch LG, Scherr PA, Wetle T.

Short-term variability of self-report of incontinence in older persons. J Am Geriatr Soc 1994; 42: 202–7.

21. Bearden DM, Allman RM, Sundarum SV, Burst BM, Bartolucci AA. Age-related variability in the use of cardiovascular imaging procedures. J Am Geriatr Soc 1993; 41: 1075–82.

22. Krauss IK. Between- and within-group comparisons in ageing research. In: Poon LW, ed. Aging in the 1980s: Psychological Issues. Washington, DC: American Psychological Association, 1980; 542–551.

23. Bornstein R, Smircina MT. The status of the empirical support for the hypothesis of increased variability in aging populations. Gerontologist 1982; 22: 258–60.

24. Hale S, Myerson J, Smith GA, Poon LW. Age, variability, and speed: between-subjects diversity. Psychol Aging 1988; 3: 407–10.

25. Greenspan SL, Rowe JW, Maitland LA, McAloon-Dyke M, Elahi D. The pituitary-adrenal glucocorticoid response is altered by gender and disease. J Gerontol 1993; 48: M72–7.

26. Ligthart GJ, Corberand JX, Fournier C *et al.* Admission criteria for immunogerontological studies in man: the SENIEUR protocol. Mech Ageing Dev 1990; 55: 89–105.

27. Greig CA, Young A, Skelton DA *et al.* Exercise studies with elderly volunteers. Age Ageing 1994; 23: 185–9.

28. Fraser CG. Age-related changes in laboratory test results. Clinical implications. Drugs Aging 1993; 3: 246–57.

29. Guay DRP, Vance-Bryan K, Gilliland S, Rodvold K, Rotschafer J. Comparison of vancomycin pharmacokinetics in hospitalized elderly and young patients using a Bayesian forecaster. J Clin Pharmacol 1993; 33: 918–22.

30. Werner P, Koroknay V, Braun J, Cohen-Mansfield J. Individualized care alternatives used in the process of removing physical restraints in the nursing home. J Am Geriatr Soc 1994; 42: 321–5.

Received 17 June 1996